

Revista do Ministério Público do Estado do Rio de Janeiro

Repositório autorizado de jurisprudência
do Egrégio Supremo Tribunal Federal
Registro nº 25/99, de 22/04/1999
DJU nº 72, de 16/04/1999, p.1

Repositório autorizado de jurisprudência
do Egrégio Superior Tribunal de Justiça
Registro nº 37 – Portaria nº 1, de 26/10/1998
DJU de 05/11/1998, p.137 - Registro retificado
Portaria nº 9, de 14/06/1999 – DJ 22/06/1999

nº 76 abr./jun. 2020

Raciocínios probabilísticos implícitos e o papel das estatísticas na análise probatória

Implicit probabilistic thinking and the role of statistics in evidentiary analysis

Edilson Vitorelli*

Sumário

1. Proposta. 2. Confiabilidade das estatísticas. 3. O probabilismo jurídico. 4. O problema: probabilidades não são intuitivas. 5. As “*naked statistics*”. 5.1. Estatísticas na comprovação de ilícitos relacionados a discriminação. 5.2. “*Naked statistics*” no Brasil. 6. As probabilidades ocultas nos raciocínios probatórios “tradicionais”. 7. Conclusão. Referências.

Resumo

Este artigo trata da utilização de estatísticas como meio de prova. Sustenta-se que os raciocínios probabilísticos são inerentes a todo tipo de prova, já que nenhuma prova é capaz de estabelecer juízos objetivos de verdade, mas apenas de elevar a probabilidade de que os fatos relevantes sejam verdadeiros. Nesse contexto, a análise tradicional, que preconiza caber ao juiz avaliar subjetivamente o quanto as provas são capazes de sustentar a afirmação da parte, desconsidera as pesquisas que demonstram que as análises de probabilidade e recorrência são altamente sujeitas a vieses cognitivos. Assim, o artigo conclui que análises estatísticas puras poderiam embasar a conclusão da ocorrência de um fato, inclusive de forma mais confiável do que a inferência subjetiva por parte do juiz.

Abstract

This paper analyses the use of statistics as evidence in civil procedure. It states that every kind of evidence has in it a probabilistic component, as no evidence is capable of defining a fact as true, but only of elevating the probability that a relevant fact is true. In this context, the traditional idea, according to which the judge should evaluate

* Pós-Doutor em Direito pela Universidade Federal da Bahia com estudos no *Max Planck Institute for Procedural Law*. Doutor em Direito pela Universidade Federal do Paraná. Mestre pela Universidade Federal de Minas Gerais. Procurador da República. *Visiting scholar* na *Stanford Law School*. *Visiting researcher* na *Harvard Law School*. Professor e coordenador de pós-graduação na Universidade Presbiteriana Mackenzie. Professor e coordenador de ensino na Escola Superior do Ministério Público da União. Professor do mestrado da Universidade Católica de Brasília.

subjectively if a piece of evidence is capable of supporting the litigator statement ignores researches that suggest that probability and recurrence analysis are highly influenced by cognitive biases. Therefore, the paper concludes that pure statistical evidence could not only support the truthfulness of a fact, but also that it could do so in a more reliable way than a subjective inference by the judge.

Palavras-chave: Meios de prova. Estatísticas. Vieses cognitivos.

Keywords: Evidence. Statistics. Cognitive biases.

1. Proposta

Este artigo trata da utilização de análises estatísticas como ferramenta para combater os vieses cognitivos do juiz, no contexto da demonstração da veracidade ou confiabilidade de uma proposição apresentada pelas partes.

Alguns exemplos da aplicação que se busca realizar podem ser apresentados, de início, a fim de ilustrar o debate que será conduzido. Em primeiro lugar, considere-se o caso do acidente de ônibus nos Estados Unidos: uma pessoa foi vítima de um acidente causado por um ônibus em determinado local. Não há qualquer prova de qual empresa seja a proprietária do veículo. Contudo, há apenas uma empresa que opera linhas regularmente naquela rua. É possível que essa empresa seja condenada a indenizar a vítima apenas em decorrência da elevada probabilidade estatística de ser a proprietária do ônibus?¹

Em um segundo exemplo, também relacionado à questão probatória, é possível considerar a situação em que diversas pessoas responderam da mesma forma às questões de uma prova para ingressar em um cargo público, cometendo os mesmos erros e obtendo os mesmos acertos. Seria possível desclassificá-las do concurso ou acusá-las de participar de uma fraude, apenas com base nessa evidência?²

Em ambos os casos, o leigo diria, intuitivamente, que a condenação é merecida. O jurista diria – e os tribunais efetivamente o disseram – que a absolvição é a providência adequada, tendo em vista que a probabilidade de que algo ocorra não quer dizer que ela ocorreu, naquele caso. Assim, na ausência de outras provas, probabilidades não bastam.

2. Confiabilidade das estatísticas

A estatística é um ramo da matemática que se incumbe de ordenar e analisar dados, a fim de permitir que eles sejam exibidos (estatística descritiva) ou que deles se extraiam inferências (estatística inferencial). Desse conceito fundamental já é possível

¹ Esse exemplo é mencionado por Susan Haack (HAACK, Susan. Evidence Matters. New York: Cambridge University Press, 2014, p. 26), e será mais bem desenvolvido ao longo do trabalho.

² O exemplo também é real, desenvolvido a partir de um caso ocorrido no estado brasileiro de Rondônia, julgado em 2011.

perceber que os resultados das análises dependerão, fundamentalmente, da qualidade do banco de dados, ou seja, da amostra analisada (sua abrangência e confiabilidade), bem como do método analítico aplicado. Há, portanto, uma dupla incerteza que cerca qualquer análise estatística: a possibilidade de que o método não seja confiável, isto é, que haja um erro humano, imputável ao analista ou que o banco de dados não seja confiável (*"garbage in, garbage out"*). Essa incerteza fomentou a popularidade de um dito alemão (falsamente atribuído a Winston Churchill, em algumas ocasiões), que assevera que "as únicas estatísticas em que se pode acreditar são aquelas que você mesmo falsificou"³. Se o banco de dados não for confiável ou se a metodologia analítica não for adequada, os resultados serão imprestáveis.

Outro problema da análise estatística diz respeito às inferências que se pode fazer a partir de seu resultado. Por exemplo, se uma pesquisa com amostra abrangente de gestores públicos demonstra que 40% deles consultam as decisões dos órgãos de controle para definir políticas públicas, disso se poderia inferir que há excesso de influência da atividade de controle sobre a gestão pública, já que se supõe que a gestão pública seja preconizada pelo que mais interessa ao público, não por aspectos procedimentais. Mas também se poderia inferir, a partir do mesmo resultado, que 60% dos gestores tomam decisões sem levar em conta o pensamento dos órgãos de controle e que, em um Estado de Direito, os fins não justificam os meios. Não se podem promover políticas corretas pelos meios errados. Nenhuma dessas inferências é falsa, a partir do resultado apresentado. O que se pode debater é se esse resultado, sozinho, é capaz de sustentar qualquer uma delas, mas isso não invalida o resultado em si.

Além de permitir a organização dos dados coletados, a partir de determinados critérios, as análises estatísticas podem ser utilizadas para permitir que as inferências sejam obtidas sem que toda a população ou todos os dados relevantes sejam coletados. É claro que, em tese, sempre é possível entrevistar todas as pessoas do país e saber exatamente, por exemplo, quantas delas têm mais de um televisor. Mas isso é caro, demorado e pouco prático. A proposta da análise estatística é obter a mesma informação por intermédio da coleta de dados de apenas uma fração relevante de toda a população. Se a amostra for representativa e a metodologia for adequada e a análise se ativer ao constructo que se pretende descrever, os resultados obtidos podem ser extrapolados para o todo⁴.

Nesse sentido, embora nem toda análise estatística lide com a definição de probabilidades, essa é uma de suas grandes utilidades. Estima-se o quão provável é o fato p , desconhecido, dada a ocorrência, já determinada, de q . Por mais estranho que isso possa parecer ao jurista, acostumado com a noção de prova como uma demonstração direta do fato, em realidade, probabilidades são aceitas cotidianamente. Elementos de prova corriqueiros, inclusive na jurisdição penal, são baseadas na probabilidade, e não

³ "Traue keiner Statistik, die du nicht selber gefälscht hast".

⁴ Sobre os conceitos fundamentais de estatística, ver WHEELAN, Charles. *Naked Statistics: Stripping the Dread from the Data*. New York: W. W. Norton & Company, 2014; FREEDMAN, David; PISANI, Robert; PURVES, Roger. *Statistics*. 4. Ed. New York: W. W. Norton & Company, 2007.

na certeza, de indicarem um fato. Assim, por exemplo, a comparação de impressões digitais, de semelhanças vocais e, até mesmo, de DNA, supõem que a probabilidade – e não a certeza – de que duas pessoas diferentes compartilhem esses caracteres, em determinado grau de similitude, é baixa demais para ser considerada.

E, no entanto, essas exceções acontecem. Em relação à comparação impressões digitais, há vasta literatura apontando possibilidades de erros, vieses cognitivos e situações de inesclarecibilidade, que parecem demonstrar que o grau de subjetividade da identificação datiloscópica é significativamente maior do que o leigo supõe⁵. Em realidade, a identificação é probabilística, não inequívoca. Talvez o caso mais notório seja o de Brandon Mayfield, cujas digitais foram identificadas pelo FBI como sendo o responsável pelo atentado terrorista na estação de trens de Madrid, em 2004. Apesar de não haver evidências de que Mayfield sequer estivesse na Espanha naquele momento, ele foi preso e processado com base nessa prova. Posteriormente, acabou inocentado⁶. Mesmo o exame de DNA tem alguma possibilidade, embora muito baixa, de falsos resultados.

Em síntese, o que se percebe é que o elemento de probabilidade, presente em análises estatísticas, também integra, subliminarmente, todas as provas tradicionalmente aceitas no processo, mesmo aquelas reputadas científicas. Quando se avançam para outros elementos de prova, como documentos ou testemunhas, esse componente probabilístico é ainda mais intenso. Supõe-se que a testemunha não mente, bem como que o documento é verdadeiro, o que pode até ser provável, à luz das circunstâncias, mas nunca será certo, ao menos não quando se toma certeza no sentido de inexistência de possibilidades de outro resultado.

A má-fama das estatísticas, portanto, talvez seja injustificada. As polêmicas decorrem das situações em que se pretende que elas desempenhem um papel mais óbvio, de maior proeminência, ou que se pretende que elas se apliquem a exemplos acadêmicos inverossímeis. Mas isso não significa que análises probabilísticas estejam ausentes em outros casos. Pelo contrário, talvez se possa dizer que toda análise probatória e, por via de consequência, toda decisão judicial que dependa de convicção de verdade quanto a fatos, implica a análise quanto à probabilidade de que determinado fato tenha ocorrido, de acordo com a prova dos autos.

⁵ Ver, por exemplo, COLE, Simon. *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Cambridge: Harvard University Press, 2001; EDMOND, Gary. *Legal versus scientific approaches to expert evidence (latent fingerprint evidence, for example): Admission, evaluation and legal system performance*. Manuscript; EDMOND, Gary. The 'Science' of Miscarriages of Justice. In: *University of New South Wales Law Journal*, vol. 37, nº 1, 2014, p. 376 e ss. É bem verdade, no entanto, que há grande disputa acadêmica acerca do número de condenações erradas decorrentes da análise de digitais. Ver, por exemplo, RISINGER, Michael. Innocents Convicted: An Empirical Justified Factual Wrongful Conviction Rate. In: *Journal of Criminal Law and Criminology*, vol. 97, 2007, p. 761 e ss.

⁶ *Mayfield v. US*, Case No. 07-35865.

3. O probabilismo jurídico

A visão mais extremada quanto ao papel que as probabilidades desempenham no processo é chamada de probabilismo jurídico (*legal probabilism*). De acordo com essa visão, seria possível definir matematicamente graus de probabilidade, que representariam os *standards* de convencimento judicial⁷. Usualmente, tais *standards* são expressos de maneira linguística, cujo conteúdo depende de concretização de acordo com as circunstâncias do caso. Na prática norte-americana, mencionam-se “além da dúvida razoável” para a condenação criminal e “causa provável” para a realização de buscas. Na esfera cível, “preponderância de provas”, “mais provável do que não”, “claro e convincente” são expressões frequentemente utilizadas⁸.

O problema é que as expressões linguísticas dos *standards* probatórios também não têm significado unívoco. Pode ser que dois juízes, apreciando o mesmo acervo probatório, em um caso criminal, divirjam acerca da presença ou não de provas além da dúvida razoável, da mesma maneira que dois juízes cíveis podem discordar acerca do cumprimento do *standard* “preponderância de provas”. Assim, quando se diz que o juiz está em dúvida, não há um parâmetro objetivo externo, pelo qual se possa verificar se essa dúvida existe efetivamente ou não⁹.

O probabilismo jurídico pretende reduzir essa incerteza por intermédio da expressão matemática dos *standards* probatórios. “Além da dúvida razoável” significaria uma probabilidade de mais de 95% de ocorrência do fato. “Causa provável” seria uma probabilidade de 30% a 51%, a depender do tribunal. “Preponderância de provas” seria um convencimento acima de 50%. Esses percentuais são, é claro, apenas uma ilustração, que serve como guia mental. Não há como medir se eles foram ou não atingidos, motivo pelo qual muitos duvidam da sua utilidade¹⁰.

Há quem, no entanto, leve mais longe a matemática. A denominada *New Evidence Scholarship*¹¹ defende a aplicação do teorema de Bayes com o objetivo de

⁷ O *standard* probatório tem, em realidade, duas funções. Ele indica, para as partes, quais provas devem produzir para obter o julgamento favorável, bem como indica, para o juiz, a quem deve ser imputado o resultado desfavorável, caso o parâmetro não seja alcançado.

⁸ Há incontáveis textos sobre o assunto. Ver, por exemplo, JAMES JR., Fleming. *Burdens of proof*. In: *Va. L. Rev.*, vol. 47, 1961, p. 51 e ss.; ALLEN, R. J. *Burdens of proof*. In: *Law, Probability and Risk*, vol. 13, nº 3-4, p. 195-219, Sept. 2014; SIMON, Rita James, and MAHAN, Linda. *Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom*. In: *Law & Society Review*, vol. 5, nº 3, 1971, p. 319-330; MCCAULIFF, C.M.A. *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees*. In: *Vand. L. Rev.*, vol. 35, 1982, p. 1239 e ss.; HAY, Bruce; SPIER, Kathryn. *Burdens of Proof in Civil Litigation: An Economic Perspective*. In: *The Journal of Legal Studies*, vol. 26, 1997 p. 413-431; PICINALI, Federico. *The burdens of proof: discriminatory power, weight of evidence, and tenacity of belief*. In: *Jurisprudence*, vol. 9, 2018, p. 192-201.

⁹ Sobre a incerteza na análise das provas, ver SHIN, Hyun Song. *The Burden of Proof in a Game of Persuasion*. In: *Journal of Economic Theory*, vol. 64, nº 1, 1994, p. 253-64. WALTON, Douglas N. *Burden of Proof*. In: *Argumentation* 2, nº 2, 1988, p. 233-54. FARLEY, Arthur M.; FREEMAN, Kathleen. *Burden of proof in legal argumentation*. In: *Proceedings of the 5th international conference on Artificial intelligence and law*, 1995, p. 156-164.

¹⁰ Para uma discussão do tema, ver LAUDAN, Larry. *Truth, Error and Criminal Law: An Essay in Legal Epistemology*. Cambridge: Cambridge University Press, 2008, especialmente o capítulo 3.

¹¹ O texto inaugural da sistematização dessa tendência parece ser LEMPERT, Richard. *The New Evidence Scholarship: Analyzing the Process of Proof*. In: TILLERS, Peter; GREEN, Eric (orgs.). *Probability and Inference in the Law of Evidence: Uses and Limits of Bayesianism*. London: Kluwer Academic Publishers, 1989, p. 62-102.

apresentar um cálculo matemático para os *standards* probatórios¹². A versão mais corrente dessa teoria é o bayesianismo subjetivo, que se vale do cálculo probabilístico como um mecanismo para orientar a definição do grau de credibilidade de uma prova, abrindo margem, no entanto, para a consideração de vieses cognitivos e valores¹³.

De modo geral, no entanto, o movimento é minoritário. A maioria dos autores concorda que não é factível converter *standards* probatórios em graus de certeza matemática e que isso pode mesmo trazer consequências práticas nefastas¹⁴. Em minuciosa análise sobre o tema, Susan Haack menciona que o probabilismo distorce o teorema matemático de Bayes, que não foi criado para essa finalidade, e que *standards* probatórios são mais bem descritos como elementos epistemológicos do que matemáticos. O juiz deverá verificar se, e até que ponto, uma prova contribui para a demonstração da ocorrência ou inoocorrência de um fato, mas não é possível “pontuar” matematicamente esse valor¹⁵.

Tomar posição nesse debate entre o probabilismo e outras correntes da epistemologia extrapolaria os limites desse trabalho. Aqui se debate apenas uma premissa: a de que probabilidades matemáticas são incompatíveis com o processo, como se elas não tivessem nenhum papel a desempenhar no universo das provas e, por isso, o único caminho possível fosse a avaliação subjetiva do juiz. É que, se não for possível definir graus de probabilidade para orientar o juiz, permitindo que sua avaliação dependa apenas de subjetividade, justificada linguisticamente pela via da argumentação, será necessário conviver com um severo contratempo, cujo risco é usualmente menosprezado: a psicologia cognitiva já demonstrou que probabilidades não são intuitivas aos seres humanos.

4. O problema: probabilidades não são intuitivas

Apesar das críticas ao probabilismo jurídico, há um problema grave, que resta sem solução, quando se descarta a hipótese de atribuir probabilidades matemáticas

Ver também PARK, Roger; SAKS, Michael. Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn. In: *Boston College Law Review*, vol. 47, 2005, p. 949–1031.

¹² O teorema de Bayes foi originalmente desenvolvido para cálculo de probabilidades em contextos não relacionados com o Direito. Ele fornece um mecanismo para se calcular a probabilidade de um evento ser verdadeiro quando se conhecem as probabilidades de outro evento.

¹³ Ver SPRENGER, Ian. The Objectivity of Subjective Bayesianism. Disponível em: <https://core.ac.uk/download/pdf/160113577.pdf>. Acesso em: 24 maio 19. Para uma crítica ao bayesianismo subjetivo, ver HAACK, Susan. Evidence Matters. New York: Cambridge University Press, 2014, p. 71.

¹⁴ Analisando o processo penal e afirmando que o julgamento matemático pode encontrar uma presunção de culpa, TRIBE, Laurence H. Trial by Mathematics: Precision and Ritual in the Legal Process. In: *Harvard Law Review*, Vol. 84, No. 6, 1971, p. 1329–1393.

¹⁵ HAACK, Susan. Evidence Matters. New York: Cambridge University Press, 2014, p. 60: “But I don’t believe it [the calculus of probabilities] could [serve as a way of computing degrees of warrant], for at least three reasons: (a) Mathematical probabilities form a continuum from 0 to 1; but because of the several determinants of evidential quality, there is no guarantee of a linear ordering of degrees of warrant. (b) The mathematical probability of (*p* and not-*p*) must add up to 1; but when there is no evidence, or only very weak evidence, eitherway, neither *p* nor not-*p* is warranted to any degree. (c) The mathematical probability of (*p* & *q*) is the product of the probability of *p* and the probability of *q*—which, unless both have a probability of 1, is always less than either; but combined evidence may warrant a claim to a higher degree than any of its components alone would do”.

aos *standards* probatórios. É que a psicologia já demonstrou que as inferências probabilísticas subjetivas estão altamente sujeitas a vieses cognitivos. Em outras palavras, probabilidades não são intuitivas¹⁶, inclusive para os juízes.

Isso significa que os juízes avaliarão o quão provável é um evento pertinente ao caso a partir de vieses de representatividade, de disponibilidade, de ancoragem e tantos outros que já foram devidamente mapeados experimentalmente. É sabido, hoje, que as pessoas tendem a achar mais prováveis eventos que sejam mais notórios, mais memoráveis ou mais assemelhados a outros que o sujeito considere familiares.

Esse problema tem sido recorrentemente subestimado pelos juristas. Amos Tversky e Daniel Kahneman demonstraram que as pessoas não apenas se equivocam ao avaliar probabilidades complexas, mas também se equivocam em testes simples, relacionados a situações corriqueiras. Informações irrelevantes influenciam na avaliação acerca do pertencimento de alguém a um grupo, apenas porque essas características tendem a estar associadas ao grupo, quer dizer, por similaridade, não por probabilidade. Esse problema, relacionado à heurística de representatividade, “leva a erros sérios, porque a similaridade ou representatividade não é influenciada por diversos fatores que deveriam afetar as avaliações de probabilidade”¹⁷. Entre esses fatores, os autores elencam a recorrente insensibilidade das pessoas aos resultados anteriores, ao tamanho da amostra e falsas correlações de causa e efeito entre eventos coincidentes.

Essas distorções cognitivas não estão limitadas a sujeitos ingênuos ou deseducados, de modo que se possa excluir a possibilidade de que afetem um juiz devidamente treinado. Pelo contrário, um experimento realizado tendo como sujeitos pesquisadores experientes demonstrou uma tendência a que eles, mesmo conhecendo as regras de metodologia científica, considerassem pequenas amostras como representativas de uma população, independentemente de qual fosse a desproporção entre uma e outra. Assim, superestimaram significativamente a confiabilidade dos resultados¹⁸.

Os trabalhos citados, já ampliados e confirmados ao longo do tempo, foram especificamente aplicados a juízes. Rachlinski e Wistrich¹⁹ listam uma série de estudos realizados por outros pesquisadores, que denotam que juízes se apoiam demais em suas intuições para decidir, o que os faz suscetíveis a erro. Esses fatores incluem a utilização de provas ou de informações não juntadas aos autos, de provas consideradas ilícitas ou a decisão com base em simples preconceito ou em reações emocionais. O que se trata aqui é de juízes que, honestamente, pretendem agir de modo imparcial, cumprindo os postulados probatórios, mas que podem ser impedidos de fazê-lo em

¹⁶ KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press, 1982.

¹⁷ KAHNEMAN, Daniel; TVERSKY, Amos. *Judgment under uncertainty: Heuristics and biases*. In: KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos. *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press, 1982, p. 4.

¹⁸ *Idem*, p. 8.

¹⁹ RACHLINSKI, Jeffrey J.; WISTRICH, Andrew J. Judging the Judiciary by the Numbers: Empirical Research on Judges. In: *Annu. Rev. Law Soc. Sci.*, 2017, p. 1-44.

virtude de limitações cognitivas inconscientes²⁰. Em um teste de resposta cognitiva (CRT), aplicado por Guthrie et al²¹ apenas a juízes, os resultados foram os mesmos que se obtêm com um grupo de adultos não treinados, quais sejam: eles erram a maioria das respostas. Não há evidência empírica de que os juízes sejam mais capazes de driblar os vieses cognitivos do que quaisquer outros indivíduos, em vários cenários²².

Isso significa que a ampliação do conhecimento e do uso das estatísticas, em quaisquer casos que se apresentem à jurisdição, seria bem-vinda. Os juízes decidiriam melhor se soubessem, estatisticamente, o quão frequentes são determinados fatos ou determinadas correlações, do que se intuissem isso a partir do que ordinariamente acontece, como determina o Código de Processo Civil Brasileiro²³ ou, em termos similares, a *Ley de Enjuiciamiento Civil* espanhola, que exige do autor prova daquilo que se desprenda do ordinário²⁴, como se os conceitos de ordinário e extraordinário fossem, de alguma forma, objetivos e não fruto de nossa percepção da realidade, recorrentemente equivocada, como demonstra a psicologia.

As evidências empíricas demonstram que os juízes, usualmente, concluem quais fatos são mais prováveis por intermédio de um *story model*. Ele propõe que a relevância das provas, o modo como elas se relacionam com outras provas e o grau de confiabilidade de cada uma delas é determinado pela criação de uma história mental que provê coerência para o conjunto²⁵. Aquilo que se encaixa na história é reputado provável, ordinário, enquanto o que não se encaixa é visto como improvável, extraordinário.

²⁰ Isso para não considerar os juízes que repetem, recorrentemente, que decidem de acordo com a sua consciência ou que gostam de olhar no olho da testemunha para saber se está mentindo.

²¹ GUTHRIE C., RACHLINSKI JJ, WISTRICH AJ. Blinking on the bench: How judges decide cases. In: Cornell Law Rev, vol. 93, 2007, p. 1-43.

²² Para uma lista mais recente de referências, ver POULTON, E C. *Behavioral Decision Theory: A New Approach*. New York: Cambridge University Press, 1994. Ver também GILOVICH, Thomas; GRIFFIN, Dale; KAHNEMAN, Daniel (eds.). *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press, 2002. GIGERENZER, G. How to Make Cognitive Illusions Disappear: Beyond “Heuristics and Biases”. *European Review of Social Psychology*, 2(1), 1991, p. 83–115; TEITELBAUM, Joshua C.; ZEILER, Kathryn (eds.). *Research Handbook on Behavioral Law and Economics*. Cheltenham: Elgar, 2018; EYAL Zamir; TEICHMAN, Doron (eds.). *The Oxford Handbook Of Behavioral Economics and the Law*. Oxford: Oxford University Press, 2014; JOLLS, Christine; SUNSTEIN, Cass; THALER, Richard. A Behavioral Approach to Law and Economics. In: *Stanford Law Review*, vol. 50, 1998, p. 1471-1550; SUSTEIN, Cass. Behavioral law and economics: a progress report. In: *American Law and Economics Review*, vol. 1, 1999, p. 115-157. Cabe ressaltar que, em algumas situações, há evidências de que os juízes são mais capazes de ignorar vieses cognitivos do que pessoas não treinadas. Em um teste aplicado por Wistrich, Guthrie e Rachlinski, relacionado ao *hindsight bias* – o viés segundo o qual a análise da probabilidade de ocorrência de um fato é influenciada pelo conhecimento de que aquele fato ocorreu – os juízes se saíram significativamente melhor do que as pessoas sem treinamento. Os juízes foram capazes de avaliar se uma busca policial foi apropriadamente conduzida, mesmo após saberem que provas incriminadoras foram encontradas. Para as pessoas comuns, o conhecimento acerca do encontro da prova cria a tendência a justificar a busca. Ver WISTRICH, Andrew; GUTHRIE, Chris; RACHLINSKI, Jeffrey. Can judges ignore inadmissible information? The difficulty of deliberating disregard. In: *University of Philadelphia Law Review*, vol. 153, 2005, p. 1251 e ss.

²³ “Art. 375. O juiz aplicará as regras de experiência comum subministradas pela observação do que ordinariamente acontece e, ainda, as regras de experiência técnica, ressalvado, quanto a estas, o exame pericial”.

²⁴ *Ley de Enjuiciamiento Civil*, Artículo 217, 2 “Corresponde al actor y al demandado reconvenirla carga de probarla certeza de los hechos de los que ordinariamente se desprenda, según las normas jurídicas a ellos aplicables, efecto jurídico correspondiente a las pretensiones de la demanda y de la reconvencción”.

²⁵ O estudo seminal quanto a essa questão foi elaborado por PANNINGTON, Nancy; HASTIE, Reid. A cognitive theory of juror decision making: the story model. In: *Cardozo Law Review*, vol. 13, 1991, p. 519

Isso implica que fatores não ligados à confiabilidade de cada prova, tais como as crenças do julgador sobre o mundo, o comportamento das pessoas e suas motivações podem exercer um papel determinante na definição de uma prova ser mais ou menos digna de crédito. Sem um guia externo, juízes poderão ser induzidos, sobretudo em casos complexos, a acreditar em provas que reforçam as suas convicções (inexatas) sobre o mundo e a desconsiderar aquelas que as contradizem. Outras questões inconscientes também podem exercer um papel inapropriado. Há evidências de que os julgadores tendem a acreditar mais em provas apresentadas em uma ordem lógica e coerente com uma história do que se as mesmas provas forem apresentadas fora de ordem²⁶. Assim, acreditar que um juiz é capaz de avaliar provas de modo puramente racional é acreditar em um mito, sem qualquer base científica. Um mito que pode até ser reconfortante, mas não é real.

5. As “*naked statistics*”

Se não é possível expressar os *standards* probatórios matematicamente e se juízos probabilísticos subjetivos são inevitáveis quando se analisam provas, parece certo que as estatísticas têm um papel a desempenhar no processo. Quando se demonstra, estatisticamente, que o consumo de um medicamento aumenta as chances de que o paciente desenvolva uma doença, o juiz pode considerar, subjetivamente, que a prova estatística contribui para o atingimento do *standard* decisório aplicável ao caso. Esse, portanto, não é o problema.

Nesse sentido, nos Estados Unidos, desde 1985, o *Manual for Complex Litigation*, editado pelo *Federal Judicial Center*, contempla pesquisas quantitativas como provas não apenas aceitáveis como desejáveis para certos tipos de demandas²⁷. Em sua edição atual, o Manual orienta os juízes no sentido de que os métodos estatísticos permitem determinar características de uma população ou universo de eventos a partir da observação dessas características em uma amostra relativamente reduzida do grupo. Assim, estudos quantitativos podem proporcionar substancial economia de tempo e de dinheiro na instrução do processo e, em alguns casos, serão “o único meio possível de coletar e apresentar dados relevantes”.²⁸

e ss. Para uma descrição mais ampla da literatura, ver EYAL, Zamir; TEICHMAN, Doron (eds.). *The Oxford Handbook Of Behavioral Economics and the Law*. Oxford: Oxford University Press, 2014, p. 518.

²⁶ PANNINGTON, Nancy; HASTIE, Reid. A cognitive theory of juror decision making: the story model. In: *Cardozo Law Review*, vol. 13, 1991, p. 519 e ss.

²⁷ Nesse sentido, durante os debates orais de *Wal-Mart Stores, Inc. v. Dukes*, et al., perante a Suprema Corte, embora relativos à admissibilidade da *classaction*, várias vezes a discussão entre os advogados e juízes derivou para questões de mérito, relativas à existência ou não de provas de uma política empresarial discriminatória e da metodologia estatística utilizada pelas partes para aferi-la ou refutá-la. O áudio dos debates está disponível em [http://www.supremecourt.gov/oral_arguments/oral_arguments.aspx]. Acesso em: 16set. 2018. É bom observar, no entanto, que algumas vezes as estatísticas são consideradas inadmissíveis, porque podem ser dúbias ou superestimadas pelo júri. Ver, por exemplo, *State v Sneed* 414 P.2d 858, 1966 e *State v Carlson*, 267 N.W. 170, 176, 1978.

²⁸ FEDERAL JUDICIAL CENTER. *Manual for complex litigation*. 4. ed. Disponível em: [<https://public.resource.org/scribd/8763868.pdf>]. Acesso em: 20 nov. 2017, p. 102.

A questão polêmica é a possibilidade de que o *standard* probatório pertinente seja atingido exclusivamente pela aplicação de probabilidades estatísticas. Wells realizou um *survey* com 740 estudantes e 111 juizes, indagando se eles estariam dispostos a impor responsabilidade ao causador de um dano, em um caso hipotético. Ambos os grupos se comportaram de modo similar. Quando o caso era apresentado tomando como base apenas provas estatísticas, o índice de condenação era inferior a 10%. Quando eram apresentadas outras provas, a condenação passava ao patamar de 65%, ainda que a probabilidade matemática e subjetiva fosse a mesma em ambos os cenários. A conclusão do autor é a de que as provas só são capazes de afetar as decisões quando contribuem para a percepção da veracidade dos fatos relevantes, não para a probabilidade de veracidade desses fatos, ainda que essa probabilidade seja muito elevada, mas inferior a 100%²⁹.

A resistência a condenações baseadas em estatísticas puras não é nova e nem se limita a estudos teóricos. Susan Haack menciona que tribunais norte-americanos tiveram a oportunidade de afirmar que as estatísticas, sozinhas, não são hábeis a sustentar o julgamento de procedência. No caso Smith, julgado em 1945, a autora pretendia responsabilizar uma empresa de ônibus por um acidente ocorrido em uma rua na qual a linha regular era operada apenas por ela. O tribunal entendeu que essa circunstância não precluiria outras possibilidades, como, por exemplo, que o acidente tivesse sido causado por um ônibus fretado. Por isso, negou o pedido³⁰.

De modo similar, no caso Sargent, julgado em 1940, uma companhia de seguros foi demandada pelo filho de uma pessoa desaparecida, com base na elevada probabilidade de que ele tivesse morrido em um acidente de caça. O tribunal negou o pedido, afirmando que “uma proposição é comprovada em um patamar de preponderância se ela aparenta ser mais provável, no sentido de uma crença concreta na sua veracidade, derivada das provas, ainda que possam existir na mente do juiz ou juizes algumas dúvidas”³¹. Ao longo do tempo, diversos outros casos tenderam a reiterar esse entendimento³².

²⁹ WELLS, Gary L. Naked Statistical Evidence of Liability: Is Subjective Probability Enough? In: *Journal of Personality and Social Psychology*, vol. 62, nº 5, 1992, p.739-752.

³⁰ Haack, op. Cit., p. 26. BETTY SMITH vs. RAPID TRANSIT INC., 317 Mass. 469, 1944.

³¹ “The weight or ponderance of evidence is its power to convince the tribunal which has the determination of the fact, of the actual truth of the proposition to be proved. After the evidence has been weighed, that proposition is proved by a preponderance of the evidence if it is made to appear more likely or probable in the sense that actual belief in its truth, derived from the evidence, exists in the mind or minds of the tribunal notwithstanding any doubts that may still linger there”. PORTER SARGENT vs. MASSACHUSETTS ACCIDENT COMPANY., 307 Mass. 246, 1940.

³² Martin Smith coleta os seguintes casos em que as estatísticas não foram admitidas como provas suficientes: Virginia & S.W. Ry. Co. v Hawk, 160 F.3d 348, 352 (6th Cir. 1908), Evans v Ely, 13 F.2d 62, 64 (3rd Cir. 1926), Sargent v Massachusetts Accident Co. 307, 246, 250, 29 N.E. 2d 825, 827 (Mass. 1940), Commercial Standards Insurance Co. v Gordon Transports Inc., 154 F.2d 390, 396 (1946), People v Collins, 438 P.2d 33 40-41 (Cal. 1968), Guenther v Armstrong Rubber Co. 406 F.2d 1315, 1318 (3d Cir. 1969), State v Carlson, 267 N.W. 170, 176 (Minn. 1978), United States v Shonubi, 103, F.3d 1085 (2d Cir. 1997), R v Watters, All ER (D) 1469 (2000). SMITH, Martin. *When Does Evidence Suffice for Conviction?* Disponível em: <https://philarchive.org/archive/SMIWDv1>. Acesso em: 15 mar. 19.

Mais recentemente, em *Sienkiewicz v. Greif*³³, a Suprema Corte do Reino Unido voltou ao assunto e discutiu longamente a relação entre estatísticas (no caso, entre epidemiologia) e decisão³⁴. Embora o caso não tenha sido decidido com base nessa premissa, quatro juízes afirmaram, em *obiterdicta*, que o nexo de causalidade entre uma doença e a exposição a determinado contaminante não poderia ser comprovado apenas por estatísticas. O ponto central dessas decisões é o de que, embora as estatísticas possam demonstrar probabilidades, elas não são capazes de determinar se o caso em exame se enquadra ou não nas probabilidades que foram estabelecidas³⁵.

5.1. Estatísticas na comprovação de ilícitos relacionados a discriminação

As estatísticas também desempenham um papel central no debate quanto a discriminação, em vários segmentos, tais como a discriminação racial ou de gênero. A demonstração da discriminação enquanto política institucional é possível apenas estatisticamente. Atos específicos, contra pessoa determinada, sempre podem ser justificados como conduta pessoal do indivíduo, não da entidade, pública ou privada, em nome do qual o indivíduo age.

Esse debate foi iniciado pela teoria do impacto desproporcional (*disparate impact doctrine*), que pretende fornecer parâmetros para a análise de condutas públicas ou privadas, aparentemente neutras, mas cuja aplicação concreta acarreta, propositadamente ou não, prejuízo maior – desproporcional – à minoria. Não se trata, assim, de investigar uma empresa que deliberadamente contrata ou promove apenas empregados brancos ou apenas homens, mas de uma empresa que adota uma política de contratação aparentemente neutra, mas com efeitos negativos sobre grupos que já estão em desvantagem social.

A teoria surgiu em 1964, no caso *Griggs v. Duke Power Co.*, o qual tem um contexto fático bastante específico. A companhia ré prestava serviços na Carolina do Norte e, no período em que a segregação racial era lícita, contratava negros apenas para o seu departamento mais subordinado (*labor department*). Com a aprovação do

³³ *Sienkiewicz v. Greif (UK) Ltd., Willmore v Knowsley Metropolitan Borough Council* [2011] UKSC10.

³⁴ O caso também gerou intensa produção acadêmica. Ver, por exemplo, STEEL, S. *Sienkiewicz v Greif and Exceptional Doctrines of Natural Causation*. In: *Journal of European Tort Law*, vol. 2, n. 3, 2011, p. 294-311; YAP, Po Jen, *Sienkiewicz v Greif: Causation and Risk of Injury*. In: *Tort Law Review*, Vol. 19, 2011, p. 61-64; LALENG, Per. *Sienkiewicz v Greif (UK) Ltd and Willmore v Knowsley Metropolitan Borough Council: A Material Contribution to uncertainty?* In: *The Modern Law Review*, vol. 74, 2011, p. 777-793.

³⁵ Nas palavras de Lord Rodgers: “Assuming that the epidemiological study is reliable, the statistics in Lord Mackay’s example would simply indicate that, if you took 100 workmen who developed dermatitis after working in the same conditions, you would expect to find that 30 developed it after having showered and 70 developed it when they had not been able to shower. So, by leading the epidemiological evidence, the only “fact” that the claimant can prove and offers to prove, on the balance of probability, is that in most cases the dermatitis would have been related to the lack of showers. So, if the judge accepts the evidence, it may legitimately satisfy him, on the balance of probability, not that the claimant’s dermatitis was caused by the lack of showers, but that, in the absence of any evidence that the claimant is atypical, it is more probable than not that his dermatitis was caused by the lack of showers. In short, the chances are that it was. Whether, in any particular case, the claimant’s dermatitis was actually caused by the lack of showers is a matter of fact – and one that remains unknown, if the only available evidence is statistical”. *Sienkiewicz v. Greif*, p. 61.

Civil Rights Act, em 1964, essa conduta passou a ser ilícita. Literalmente, no dia seguinte à aprovação do ato, a empresa passou a adotar um teste escrito para a admissão e promoção de empregados, além da exigência de diploma de ensino médio. Como os negros tinham, até então, estudado em escolas segregadas de pior qualidade, a aplicação dessa norma, aparentemente neutra, os atingia de modo muito mais grave, o que foi demonstrado estatisticamente.

A Suprema Corte dos Estados Unidos decidiu que os novos requisitos de contratação impactavam desproporcionalmente sobre grupos étnicos minoritários. Nesse caso, seria ônus do empregador demonstrar que eles são razoavelmente necessários para o desempenho da função. Como disse o *Chief Justice Burger*, o *Civil Rights Act* “proscree não apenas atos de discriminação aberta, mas também práticas que são legítimas na forma, mas discriminatórias na sua operação”³⁶. De modo similar, em *Wal-Mart Stores, Inc. v. Dukes*³⁷, estudos estatísticos foram utilizados para demonstrar que mulheres não tinham as mesmas oportunidades que homens em promoções para cargos de gerência na companhia e que, se isso ocorria em todo o país, não poderia ser justificado pelas condutas individuais dos superiores hierárquicos de cada uma das vítimas, mas por uma cultura empresarial disseminada. Como se percebe, a única forma de se fazer a passagem entre a negativa de oportunidades para algumas mulheres e a comprovação de uma cultura empresarial de discriminação, sem avaliar as carreiras de cada uma das mulheres do maior empregador dos Estados Unidos, é por intermédio de estatísticas.

Apesar do sucesso inicial de *Griggs v. Duke Power Co.* e da inviabilidade de se produzir provas de discriminação disseminada de outro modo, o sucesso das estatísticas, também nesse segmento, é limitado. Em *Washington v. Davis*³⁸, caso julgado 12 anos depois, a Suprema Corte refluíu da teoria do impacto desproporcional, para exigir que, além do resultado discriminatório, os interessados demonstrassem a intenção, por parte do contratante, de praticar discriminação. Embora os contextos dos casos sejam distintos³⁹, o papel das estatísticas, pelo menos isoladamente, foi diminuído.

Posteriormente, em *Ricci v. DeStefano*⁴⁰, a importância das estatísticas viria a ser ressaltada, mas para a finalidade oposta. O corpo de bombeiros de New Haven anulou um teste de promoção ao fundamento de que nenhum bombeiro negro havia sido aprovado e isso seria evidência de que o teste causara um impacto desproporcional sobre os negros. 20 bombeiros aprovados ingressaram em juízo, alegando que a anulação da prova era discriminatória. A Suprema Corte deu-lhes

³⁶ *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).

³⁷ *Wal-Mart Stores, Inc. v. Dukes*, et al., 564 U.S. 338 (2011).

³⁸ *Washington v. Davis*, 426 US 229 (1976).

³⁹ “Disproportionate impact is not irrelevant, but it is not the sole touchstone of an invidious racial discrimination forbidden by the Constitution”. Embora não seja relevante para os fins deste artigo, não se pode dizer que o precedente de Davis superou o de Griggs, porque este analisou a conduta discriminatória sob o pálio do *Civil Rights Act*, enquanto aquele tratou diretamente da cláusula constitucional da *equal protection*.

⁴⁰ *Ricci v. DeStefano*, 557 U.S. 557 (2009).

razão, ao fundamento de que o município não tinha “firmes bases probatórias”⁴¹ de que ele teria sido responsabilizado por impacto desproporcional se tivesse promovido os bombeiros aprovados. A decisão foi fortemente criticada do ponto de vista do seu efeito negativo sobre os empregadores que queiram promover políticas de igualdade.

5.2. “Naked statistics” no Brasil

No Brasil, a situação não é distinta da que foi apresentada no tópico anterior. O Poder Judiciário já foi confrontado com a análise de estatísticas como provas em mais de uma ocasião. Em 2011, investigação da Polícia Federal, denominada “Operação Q.I.” detectou evidências estatísticas de que 17 candidatos a um cargo público no estado de Rondônia teriam fraudado o exame. Todos eles realizaram provas na mesma localidade, tiveram notas desproporcionalmente altas em relação aos concorrentes e suas respostas certas e erradas seguiam um padrão, que era distinto dos demais candidatos. Os peritos que analisaram o caso concluíram que as chances de que esse padrão fosse resultado de coincidência seriam de 1:55.313.555.376.598.900, o que é milhares de vezes menos do que as chances de alguém ganhar em qualquer loteria.

Com base nessa prova, foi proposta uma ação para retirar os candidatos do cargo, que tramitou sob o número 0018584-46.2011.8.22.0001, na Justiça de Rondônia. Embora eles tenham sido provisoriamente afastados, o pedido acabou sendo julgado improcedente por falta de provas. O juiz do caso afirmou:

A análise Estatística Matemática é uma ferramenta de estudo descritivo e comparativo de dados, ela nos apresenta um caminho, uma direção de pensamento, e nos ajuda a obter uma conclusão clara de um determinado fato. Se for necessário ela serve de apoio a outros estudos para obter resultados justos e esclarecedores. (...) a prova pericial neste caso possui valor de indícios e, portanto insuficiente ao fim que se propõe, nulidade do ato de posse ao fundamento de fraude do concurso. É forçoso concluir pela insuficiência de prova a afirmar ocorrência da suposta fraude a partir dos elementos dos autos, logo não tenho pela ocorrência de fraude a partir de laudo estatístico, pois trata-se de prova de natureza indiciária.

No mesmo sentido, o Superior Tribunal de Justiça já decidiu, em pelo menos três casos, que “a mera probabilidade, constituída a partir de laudo estatístico”⁴², não basta para comprovar fraude em concursos públicos. No precedente que inaugura a

⁴¹ Esse parâmetro, denominado “*Strong-basis-in-evidence standard*”, havia sido firmado em *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469 (1989).

⁴² Essa frase está na decisão do AgRg no RMS 26.011/DF, Rel. Ministro REYNALDO SOARES DA FONSECA, QUINTA TURMA, julgado em 30/06/2015, DJe 03/08/2015.

série⁴³, o tribunal apenas afirma que o laudo estatístico deveria ser corroborado por outros “elementos concretos (...) por estar fundado em probabilidades”. A decisão do STJ cita, ainda, trechos do acórdão recorrido que analisou a qualidade dos elementos estatísticos produzidos no processo, para impugnar-lhes a qualidade. Ocorre que a desconstrução das estatísticas, pelos julgadores que atuaram nas instâncias inferiores, foi feita de modo leigo⁴⁴, sem recurso a outro *expert* que se opusesse às conclusões que sustentavam que as chances de o resultado decorrer de coincidência eram irrisórias. O STJ não debateu essa falha, limitando-se a transcrever as objeções do acórdão recorrido. O Tribunal Superior ainda transcreve outra decisão, oriunda do Tribunal de Justiça do Distrito Federal, na qual o Desembargador Leocir Manoel da Luz aponta que, “no campo das estatísticas, a ocorrência de provas idênticas entre os aprovados no concurso era muito mais rara que acertar sozinho na megasena”. Apesar dessa baixíssima probabilidade de coincidência, o julgador, logo a seguir, complementa: “Portanto, não descartou totalmente a ocorrência de uma coincidência, quanto mais uma coincidência parcial”. Quer dizer, a decisão, citada pelo STJ, reconhece que as possibilidades de que o fato evidenciado pela análise estatística não ter ocorrido eram mínimas, irrisórias, mas concluiu que, conquanto mínimas, elas existem e, portanto, sustentam a improcedência.

Daí em diante, os dois outros casos que o STJ julgou, acerca do problema das estatísticas, se limitam a citar a primeira decisão como argumento de autoridade, para afirmar que estatísticas, sozinhas, não bastam para sustentar uma conclusão, qualquer que seja ela. Não há exame quanto a qualidade dos estudos estatísticos produzidos em cada um dos casos. Conclui-se que outras provas se fazem necessárias, além das probabilidades. O laudo estatístico seria “mera prova indiciária”⁴⁵. Parece haver, portanto, uma injustificada tariffação da prova estatística no STJ, para desvalorizá-la.

Wasserman explica essa resistência às estatísticas, apontando que elas geram a percepção no julgador de injustiça para com o réu, já que o julgamento não seria propriamente de seu caso, mas de sua própria conduta pretérita, ou do comportamento do grupo ao qual ele pertence⁴⁶. Mas o que se percebe de todos esses casos é que, de maneira geral, não há um questionamento acerca da confiabilidade das estatísticas apresentadas aos tribunais. O que se disputa é a possibilidade de inferir causalidade

⁴³ RMS 24.503/DF, Rel. Ministro ARNALDO ESTEVES LIMA, QUINTA TURMA, julgado em 15/12/2009, DJe 01/02/2010.

⁴⁴ O Desembargador João Mariosi, do Tribunal de Justiça do Distrito Federal, citado pelo STJ, afirmou: “Quando decido, não estou dando aula de Estatística, não estou dando aula de Direito, estou decidindo, sou o Estado em uma forma reservada aqui, um por dezessete, e a validade é justamente o Estado falando e decidindo a matéria. Portanto, quando se trata de prova indiciária em que se busca a Estatística, está havendo um sofisma e sério, porque o cluster, que significa caixa, usado na Estatística, foi usado indevidamente. Seria o caso de dizer que qualquer pessoa que jogou na lota seja criminoso, porque o criminoso que matou fulano de tal deixou um bilhete da lota. É o que foi feito neste processo”.

⁴⁵ A expressão é utilizada na decisão do REsp 1307532/RJ, Rel. Ministro MAURO CAMPBELL MARQUES, SEGUNDA TURMA, julgado em 09/04/2013, DJe 16/04/2013.

⁴⁶ WASSERMAN, David. The Morality of Statistical Proof and the Risk of Mistaken Liability. In: Cardozo L. Rev., vol. 13, 1992, p. 935 e ss. De modo geral, ver WASSERMAN, Larry. *All of Statistics: A Concise Course in Statistical Inference*. New York: Pittsburgh, Springer, 2003.

das probabilidades estatísticas, mesmo quando elas são elevadas. De modo simples, afirmar que “há 99% de chances de que um ônibus pertença à empresa X” parece soar a um juiz como “não há fatos que demonstrem que o ônibus pertença à empresa X, ainda que isso seja muito provável”.

Essa conclusão ignora que todas as provas, sem exceção, envolvem um ou mais juízos probabilísticos implícitos. O único motivo pelo qual essas provas são vistas como mais confiáveis é exatamente porque os juízos probabilísticos são implícitos, não explícitos. Mas elas não apenas são tão probabilísticas quanto as estatísticas puras, quanto podem conter probabilidades menores de corresponder à realidade e, por isso, ser menos confiáveis.

6. As probabilidades ocultas nos raciocínios probatórios “tradicionais”

Suponha-se que não seja justo condenar uma empresa a indenizar alguém em virtude de um câncer decorrente da exposição a um produto por ela produzido porque, estatisticamente, há apenas 80% de chances de que o produto seja a causa da doença, o que significa que há 20% de chances de que não seja. Por outro lado, suponha que seria justo condenar alguém com base no testemunho de uma pessoa quanto a um fato que ela presenciou porque o testemunho, diferentemente das estatísticas, se dirige diretamente àquele fato, não a uma classe (população) de fatos.

O problema é que esse raciocínio ignora que a decisão fundada na prova testemunhal teria que pressupor a) que há 100% de chances de que a testemunha esteja dizendo a verdade; b) que há 100% de chances de que ela tenha percebido corretamente a situação; c) que há 100% de chances de que ela tenha relatado apropriadamente a situação em juízo. Se, como parece evidente, for possível afirmar que qualquer desses três fatores não atinge o patamar de 100%, então se estará diante de uma probabilidade de que o testemunho seja verdadeiro e capaz de demonstrar a ocorrência daquele fato. Se, por exemplo, cada um desses fatores tiver probabilidade de 90% e, partindo do pressuposto de que eles estejam relacionados, a probabilidade final de que a decisão esteja fundada em um fato verídico será de apenas 73%.

Esse raciocínio pode ser expandido para qualquer tipo de prova, em qualquer situação. Tome-se como ilustração o caso do testemunho ocular (*Eye witness testimony*), que é recorrentemente considerado uma prova robusta, eis que decorre do conhecimento direto do fato pela testemunha, que o presenciou. Apesar de as pessoas serem condenadas civil e criminalmente, todos os dias, em decorrência de testemunhos oculares, há diversos estudos apontando a baixa confiabilidade desse tipo de prova⁴⁷. Em um *survey* com 63 especialistas em testemunho ocular,

⁴⁷ O primeiro estudo de que se tem notícia sobre o assunto é de MÜNSTERBERG, Hugo. *Onthewitness stand*. New York: Doubleday, Page & Company, 1908. Esse trabalho recebeu um interessante comentário, por ocasião de seu centenário, em MEMON, Amina; MASTROBERARDINO, Serena; FRASER, Joanne. Münsterberg's Legacy: What Does Eyewitness Research Tell Us About the Reliability of Eyewitness Testimony? In: *Appl. Cognit. Psychol.* Vol. 22, 2008, p. 841–851. Ver também MAGNUSSEN, S.; MELINDER, A.; STRIDBECK, U.; RAJA, A. Q. Beliefs about factors affecting the Reliability of eyewitness testimony: A Comparison of judges, jurors

Kassin, Ellsworth e Smith⁴⁸ encontraram elevado grau de concordância (80%) de que esse tipo de prova pode ser negativamente influenciado não por três, mas por pelo menos 9 fatores: o modo de fazer a pergunta à testemunha, as instruções durante a identificação da pessoa visualizada, informações enganosas recebidas após o fato, a correlação entre precisão e confiabilidade, atitudes e expectativas, o tempo de exposição, transferência inconsciente, identificação cara-a-cara (“showups”) e a curva de esquecimento. Além disso, um número significativo de especialistas (70%) ainda acrescenta preconceitos raciais, diferenças no posicionamento das pessoas a serem identificadas e a tendência a superestimar a duração dos eventos, o que eleva o número de fatores de dúvida a 12⁴⁹.

Aplicando o mesmo raciocínio desenvolvido acima, se for possível afirmar que cada um desses marcadores de influência negativa reduz a credibilidade de uma testemunha ocular em apenas 10%, e supondo que, na ausência de qualquer deles, a testemunha seria 100% crível, seria preciso concluir que 80% dos especialistas entrevistados por Kassin, Ellsworth e Smith concordariam que o grau máximo de credibilidade dessa testemunha ocular seria de 38,7%. Se o escrutínio for elevado aos 12 critérios, preconizados por 70% dos experts, uma testemunha ocular teria apenas 28,2% de credibilidade, mantido o fator de dedução de 10% por critério.

Esse raciocínio poderia ser expandido para qualquer tipo de prova, com maior ou menor grau de credibilidade. Documentos públicos ou particulares, testemunhos, perícias e quaisquer outras provas apenas são hábeis a sustentar determinada afirmação na medida em que estiverem isentos de dúvidas. Como essa medida

and the general public. *Appl. Cognit. Psychol.*, vol. 24, 2010, p. 122-133; WELLS, Gary; OLSON, Elizabeth. Eyewitness Testimony. In: *Annual Review of Psychology*, Vol. 54, 2003, p. 277-295; LEACH, Amy-May; CUTLER, Brian L.; VAN WALLENDAL, Lori. Lineups and Eyewitness Identification. In: *Annual Review of Law and Social Science*, Vol. 5, 2009, p. 157-178; CUTLER, Brian L.; PENROD, Steven; MARTENS, Todd. The reliability of eyewitness identification. In: *Law and Human Behavior*, vol. 11, n. 3, 1987, p. 233-258; BARKOWITZ, P.; BRIGHAM, J. C. Recognition of faces: Own-race bias, incentive, and time delay. In: *Journal of Applied Social Psychology*, vol. 12, 1982, p. 255-268; BRIGHAM, J. C. The accuracy of eyewitness evidence: How do attorneys see it? In: *The Florida Bar Journal*, vol. 55, 1981, p. 714-721; FRAZZINI, S. F. Review of eyewitness testimony. In: *The Yale Review*, vol. 70, 1981, p. XVIII-XX; LOFTUS, E. F. *Eyewitness testimony*. Cambridge, Massachusetts: Harvard University Press, 1979.

⁴⁸ KASSIN, S. M., ELLSWORTH, P. C., SMITH, V. L. The “general acceptance” of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, vol. 44, n. 8, 1989, p. 1089-1098.

⁴⁹ Martin Smith faz uma consideração similar à desenvolvida no texto, embora ele não subscreva a conclusão: “For a court surely would – and surely should – find the Blue-Bus company liable on the basis of other kinds of defeasible evidence. Suppose that, instead of the market share evidence, an eyewitness to the incident testifies that the bus involved was a Blue-Bus bus. Provided this testimony is not contradicted or called into question, it would generally be deemed sufficient for a finding of liability. But eyewitnesses are fallible. The fact that a witness testifies that the bus involved was a Blue-Bus bus does not guarantee that it was. The witness might have suffered a hallucination, or another sort of bus might have been sporting a Blue-Bus logo, or the witness might have been simply lying in order to smear the company etc. If we were forced to come up with some numerical estimate of how likely it is that the bus really was a Blue-Bus bus, given the witness testimony, it’s doubtful that we would go quite as high as 90% – that would seem overly trusting. But 90% is, of course, precisely how likely it is that the bus involved was a Blue-Bus bus, given the statistical evidence about which we were so apprehensive. A verdict of liability based on this testimonial evidence would actually run a greater risk of error than a verdict of liability based upon the statistical evidence. And yet the former verdict would seem acceptable, while the latter would not”. SMITH, Martin. When Does Evidence Suffice for Conviction? Disponível em: <https://philarchive.org/archive/SMIWDv1>. Acesso em: 15 mar. 19.

nunca será de 100%, sempre se poderá dizer que determinada prova, ou conjunto de provas, sustentam a probabilidade de uma conclusão, em maior ou menor grau. Exatamente como as estatísticas.

Do mesmo modo, não é possível reduzir o valor das provas estatísticas com o argumento de que elas se debruçam sobre o perfil de uma população, não sobre o caso em si e, portanto, o réu seria condenado por pertencer a determinado grupo, não pelo que efetivamente fez. Sempre que o juiz afere em que grau as provas do autor e do réu são críveis, é necessário generalizar o grau de credibilidade de cada uma das provas apresentadas, assim como do conjunto que elas formam. Isso implica avaliar até que ponto é provável que uma testemunha esteja dizendo a verdade, até que ponto é provável que determinado método pericial seja confiável, até que ponto o depoimento das partes aparenta ser verídico etc. E, embora os sistemas jurídicos ocidentais permitam que os juízes valorem essas ocorrências à luz das circunstâncias dos casos, há abundantes estudos psicológicos que demonstram que essa análise está inerentemente sujeita a vieses cognitivos.

Dito de outro modo, quando um juiz avalia se uma testemunha é mais digna de crédito que outra, ele está avaliando até que ponto é provável, à luz do conjunto probatório do caso, que aquela pessoa esteja dizendo a verdade. Isso é específico do caso e, via de regra, é o que vai constar explicitamente da decisão. Porém, implicitamente, o juiz também analisa:

A) até que ponto provas testemunhais são confiáveis ou não, naquele caso específico. Por exemplo, se o litígio for referente à execução de um contrato de obra pública para a construção de uma ponte, o juiz provavelmente reputará a prova testemunhal menos confiável que o diário de obra. Isso não significa que a prova testemunhal, naquele caso, seja, de fato, menos confiável, mas apenas que, por generalização, a população de provas testemunhais costuma ser menos confiável em casos nos quais há abundância de elementos documentais.

B) até que ponto provas testemunhais são críveis naquele tipo de litígio, de maneira geral. Por exemplo, em litígios relacionados à execução de quaisquer contratos públicos, provas testemunhais são pouco críveis, porque a administração pública não procede, de modo geral, de maneira verbal. “Seria estranho”, avaliaria o juiz, se apenas nesse caso o testemunho fosse mais verídico que o diário de obra já que, via de regra, diários de obras são mais confiáveis que testemunhas. Essa redução da credibilidade do testemunho não depende do caso em si, mas da opinião que o juiz tem quanto àquela população de casos. E pode ser que, naquele caso, a testemunha seja mesmo mais confiável que o diário de obra.

C) Até que ponto um testemunho de qualquer pessoa naquela situação seria crível. Por exemplo, testemunhos de ex-empregados da empresa, que com ela tenham outros litígios, são inerentemente considerados menos críveis, mesmo que não se saiba se, naquele caso, essa circunstância exerce ou não influência. O juiz será levado a crer que, geralmente, testemunhas que estão litigando contra a parte, ainda que por motivos completamente distintos do caso em tela, merecem menos confiança, pois podem guardar algum tipo de rancor. Perceba-se que esse raciocínio decorre apenas de uma generalização quanto ao comportamento de uma população de testemunhas que conduzem processos contra as partes dos casos nos quais testemunham, não da análise de evidências de que, naquele caso, aquela testemunha guarda algum rancor.

D) Até que ponto aquele meio de prova é crível, em geral. Embora os sistemas ocidentais tenham abandonado, como regra, a tarifação prévia de provas, é inegável que determinados meios são considerados inerentemente mais confiáveis que outros, mesmo que não haja evidências de que essa confiabilidade seja justificável, naquele caso. A confissão é usualmente reputada a mais crível das provas, seguida do documento público, do documento particular, do testemunho direto e do testemunho indireto (*hearsay*).

Em resumo, todo raciocínio probatório inclui análises probabilísticas implícitas de credibilidade que, se fossem explicitadas, teriam potencial para reduzir a credibilidade das provas aquém dos índices que se obtém em algumas situações em que análises estatísticas puras estão em jogo. Além disso, muitas vezes o raciocínio probatório inclui generalizações implícitas quanto a um conjunto de situações que envolvem aquela prova, não relacionados ao caso em exame. Da mesma forma que estatísticas dependem da população analisada, não do caso singular, a credibilidade das provas e, com isso, a decisão também depende de generalizações a respeito da população de circunstâncias que cercam aquele caso, ou aquele tipo de caso, não apenas do caso em si e das provas que dizem respeito diretamente a ele.

Logo, excluir a possibilidade de que alguém seja condenado apenas com base em estatísticas significa, primeiramente, tarifar negativamente a prova estatística, eis que se aceita, sem maiores dificuldades, que alguém seja condenado apenas com base em um documento (independentemente da possibilidade de ele ser falso, ainda que pareça verdadeiro), ou apenas em uma testemunha (independentemente da possibilidade de que ela possa estar mentindo, ainda que pareça dizer a verdade) ou com base em uma análise pericial (independentemente da possibilidade de falha do método científico). A única diferença é que a estatística traz em si, explicitamente, o seu grau de confiabilidade enquanto, nos demais tipos de prova, essa possibilidade de erro permanece implícita, oculta, ao aguardo de que surjam outras evidências que a revelem.

Em segundo lugar, significa admitir que alguns réus não serão condenados, mesmo que haja elevada probabilidade de serem culpados, enquanto outros serão condenados, mesmo que haja maiores probabilidades do que os primeiros de que sejam inocentes. No exemplo das fraudes em concursos públicos, apresentado acima, há pessoas que efetivamente foram eliminadas dos exames por portarem objetos proibidos, mesmo que não houvesse qualquer evidência de que o tivessem utilizado para obter benefícios e, pelo contrário, fosse baixa a probabilidade de que o objeto tivesse potencial para influir no resultado. Por outro lado, pessoas cujas probabilidades de não terem cometido fraude eram irrisórias foram absolvidas, apenas porque a única coisa que pesava contra elas não era uma prova “tradicional”, mas apenas a elevada probabilidade de serem fraudadores.

Em terceiro lugar, significa admitir que alguns réus jamais serão condenados. No mercado consumidor, isso implica aceitar que determinados seguimentos de mercado serão proporcionalmente super-responsabilizados (*overdeterrence*), enquanto outros serão sub-responsabilizados (*underdeterrence*). Aquelas atividades cujos efeitos deletérios sejam facilmente perceptíveis serão submetidas a um patamar de responsabilidade civil superior às suas concorrentes que forem capazes de mascarar os sinais visíveis dos danos que causam. A longo prazo, isso significa que todo o mercado será estimulado a direcionar seus esforços para a ocultação dos danos, evitando aquilo que puder ser provado de acordo com métodos “tradicionais”, em vez de direcioná-los para a redução concreta dos riscos mais significativos. De modo similar, excluir a possibilidade de condenações por estatísticas significa dizer que empresas que disseminam uma cultura de condutas lesivas a minorias, raciais ou de gênero, permanecerão impunes, eis que poderão sempre justificar-se pelo argumento de que se trata de “ato isolado”.

Apesar de tudo, é bom anotar que as decisões baseadas apenas em provas estatísticas continuam sendo amplamente minoritárias e, de modo geral, a técnica é repudiada nos tribunais, salvo como elemento acessório a outras provas. Talvez o mais conhecido dos casos de processos relacionados a demonstrações estatísticas seja *McCleskey v. Kemp*,⁵⁰ no qual a tese de um condenado à morte, para demonstrar a injustiça da punição, era a de que um negro que cometesse um crime contra uma vítima branca, no estado da Geórgia, tinha 4,3 vezes mais chances de ser condenado à morte do que outros homicidas. A conclusão foi obtida a partir de um estudo acadêmico que analisou mais de 500 casos e controlou 39 variáveis.⁵¹ Contudo, em um resultado de cinco votos contra quatro, McCleskey teve seu pedido rejeitado e foi executado em setembro de 1991.⁵²

⁵⁰ 481 U.S. 279 (1987).

⁵¹ O estudo foi publicado anos antes de ser utilizado no caso. Trata-se de BALDUS, David C.; PULASKI, Charles; WOODWORTH, George. Comparative review of death sentences: an empirical study of the Georgia experience. *The Journal of Law & Criminology*, vol. 74, n. 3, p. 661-753, 1983.

⁵² Como se observa, a evidência estatística foi suficiente para convencer 4 juízes. A decisão é recorrentemente citada no rol das piores adotadas pela Suprema Corte em todos os tempos.

7. Conclusão

Do que foi exposto até aqui, é possível concluir que a afirmação, *ex-ante*, de que réus não podem ser condenados apenas com base em apurações estatísticas decorre de uma má compreensão do papel que as probabilidades desempenham no raciocínio probatório, de maneira geral. Demonstrou-se que qualquer prova implica raciocínios probabilísticos quanto a sua confiabilidade, os quais dependem de raciocínios probabilísticos quanto à confiabilidade daquela espécie de prova, de modo geral. Disso decorre que uma prova nunca é passível de ser analisada apenas de acordo com o caso, mas exige raciocínios probabilísticos implícitos, que não se distinguem dos raciocínios estatísticos.

Em segundo lugar, a impossibilidade de se condenar alguém com base em estatísticas se baseia na falsa premissa de que elas se dirigem ao comportamento geral de um grupo, enquanto as provas tradicionais se dirigem diretamente ao caso. Em realidade, graças ao raciocínio probabilístico antes demonstrado, qualquer prova, mesmo que se dirija diretamente ao caso, tem apenas um percentual (usualmente implícito e não calculado) de chances de, efetivamente, comprovar as afirmações da parte. Dito de modo diverso, as provas tradicionais de responsabilidade por determinado ato só aparentemente comprovam ou não a sua ocorrência. Como elas sempre podem ser verdadeiras ou falsas, há apenas uma probabilidade de que o réu seja culpado ou inocente, assim como as estatísticas demonstram um grau de probabilidade de que tal pessoa seja responsável por um ato. Ocorre que, em estatísticas, a margem de erro é calculada e explicitada. Nas provas tradicionais, isso não ocorre.

Em terceiro lugar, o desestímulo ao uso de estatísticas no âmbito probatório se torna ainda mais nefasto quando a lei concita o juiz a avaliar as provas produzidas “de acordo com o que usualmente acontece”, como se fosse possível que o juiz tivesse uma visão imparcial acerca do que “usualmente acontece” ou que “se desprende do ordinário”. A psicologia comportamental já estabeleceu que, em decorrência dos vieses cognitivos, é impossível que um ser humano avalie, de modo objetivo, o que é ordinário ou extraordinário, corriqueiro ou excepcional. Pelo contrário, probabilidades não são atributos que o cérebro humano seja capaz de perceber e processar instintivamente. Essa análise é influenciada por elementos que nada têm a ver com a efetiva frequência de ocorrência ou recorrência de um fato, o que pode guiar o juiz no sentido de uma avaliação imperfeita das provas.

Em quarto lugar, a exclusão das estatísticas como meios de prova fornece estímulos negativos ao mercado, fomentando comportamentos negativos à sociedade, consubstanciados em condutas que não possam ser comprovados por evidências diretas, tradicionais. A sociedade nada tem a ganhar com a exclusão de meios de prova ou com a inocência de réus cujas probabilidades de serem culpados é elevada, apenas porque essas probabilidades decorrem unicamente de estatísticas.

Nesse contexto, os sistemas jurídicos ocidentais teriam a ganhar se investissem na regulamentação da prova estatística, em vez de insistirem na sua exclusão. Seria melhor fomentar um diálogo produtivo entre juristas e estatísticos para definir, por

exemplo, a) quais tipos de direitos podem ser definidos por estatísticas; b) quais métodos estatísticos são mais ou menos confiáveis em quais tipos de casos; c) qual o grau de probabilidade exigível para cada caso; d) de que modo a prova estatística deve ser produzida; e) de que maneira as partes podem intervir na análise realizada pelo perito, caso ele seja designado pelo juiz; f) quais as credenciais a serem exigidas do *expert* designado para realizar a análise; g) como os juízes devem analisar as estatísticas para tomar suas decisões e assim por diante⁵³. Nesse contexto, os Códigos Modelos para a Ibero-América poderiam ser aperfeiçoados por normas que regulassem, de modo mais minudente, o uso das estatísticas.

O receio de que as estatísticas possam levar a uma condenação equivocada torna inesclarecíveis questões que poderiam ser decididas com acuidade. Via de regra, isso desfavorecerá a parte autora, que terá que conviver com a perda do direito material. Por alguma razão, os juristas e o sistema judicial parecem mais confortáveis em admitir que o autor fique sem o direito material que deveria ter do que em prejudicar o réu pelo uso de uma prova que poderia vir a ser inadequada. Esse é um modo equivocado de se encarar a realidade. A omissão em corrigir um mal deveria pesar tanto sobre a consciência do juiz quanto o risco de corrigi-lo equivocadamente.

Referências

- AFSHARTOUS, David, Sample Size Determination for Binomial Proportion Confidence Intervals: An Alternative Perspective Motivated by a Legal Case. In: *The American Statistician*, vol. 62, 2007, p. 27 e ss.
- AITKEN, C. C. G.; TARONI, Franco. *Statistics and the Evaluation of Evidence for Forensic Scientists*. 2. Ed. New Jersey: Wiley, 2004.
- ALLEN, R. J. Burdens of proof. In: *Law, Probability and Risk*, vol. 13, nº 3-4, 2014, p. 195-219.
- BALDUS, David C.; PULASKI, Charles; WOODWORTH, George. Comparative review of death sentences: an empirical study of the Georgia experience. In: *The Journal of Law & Criminology*, vol. 74, nº 3, p. 661-753, 1983.
- BARNES, David W., *Statistics as Proof: Fundamentals of Quantitative Evidence*. Boston: Little, Brown and Company, 1983.

⁵³ No sentido do texto, discutindo metodologias estatísticas e sua aplicabilidade aos processos judiciais, ver, por exemplo, AFSHARTOUS, David, "Sample Size Determination for Binomial Proportion Confidence Intervals: An Alternative Perspective Motivated by a Legal Case". In: *The American Statistician*, vol. 62, 2007, p. 27 e ss.; AITKEN, C. C. G.; TARONI, Franco. *Statistics and the Evaluation of Evidence for Forensic Scientists*. 2. Ed. New Jersey: Wiley, 2004; BARNES, David W. *Statistics as Proof: Fundamentals of Quantitative Evidence*. Boston: Little, Brown and Company, 1983; BARNES, David W.; CONLEY, John M. *Statistical Evidence in Litigation: Methodology, Procedure, and Practice*. Boston: Little, Brown and Company, 1986; DOWNTON, F. Legal Probability and Statistics. In: *145 Journal of the Royal Statistical Society*, vol. 145, 1982, p. 395 e ss.; FIENBERG, Stephen E.; STRAF, Miron. Statistical Evidence in the US Courts: An Appraisal. In: *Journal of the Forensic Science Society*, vol. 31, 1991, p. 259 e ss., entre diversos outros.

_____; CONLEY, John M. *Statistical Evidence in Litigation: Methodology, Procedure, and Practice*. Boston: Little, Brown and Company, 1986.

BRIGHAM, J. C. The accuracy of eyewitness evidence: How do attorneys see it? In: *The Florida Bar Journal*, vol. 55, 1981, p. 714–721.

COLE, Simon. *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Cambridge: Harvard University Press, 2001.

CUTLER, Brian L.; PENROD, Steven; MARTENS, Todd. The reliability of eyewitness identification. In: *Law and Human Behavior*, vol. 11, nº 3, 1987, p. 233–258.

DOWNTON, F. Legal Probability and Statistics. In: *145 Journal of the Royal Statistical Society*, vol. 145, 1982, p. 395 e ss.

EDMOND, Gary. *Legal versus scientific approaches to expert evidence (latent fingerprint evidence, for example): Admission, evaluation and legal system performance*. Manuscrito.

_____. The 'Science' of Miscarriages of Justice. In: *University of New South Wales Law Journal*, vol. 37, nº 1, 2014, p. 376 e ss.

EYAL Zamir; TEICHMAN, Doron (eds.). *The Oxford Handbook Of Behavioral Economics and the Law*. Oxford: Oxford University Press, 2014.

FARLEY, Arthur M.; FREEMAN, Kathleen. Burden of proof in legal argumentation. In: *Proceedings of the 5th international conference on Artificial intelligence and law*, 1995, p. 156-164.

FIENBERG, Stephen E.; STRAF, Miron. Statistical Evidence in the US Courts: An Appraisal. In: *Journal of the Forensic Science Society*, vol. 31, 1991, p. 259 e ss.

FRAZZINI, S. F. Review of eyewitness testimony. In: *The Yale Review*, vol. 70, 1981, p. XVIII-XX.

FREEDMAN, David; PISANI, Robert; PURVES, Roges. *Statistics*. 4. Ed. New York: W. W. Norton & Company, 2007.

GIGERENZER, G. How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases." In: *European Review of Social Psychology*, 2(1), 1991, p. 83–115.

GILOVICH, Thomas; GRIFFIN, Dale; KAHNEMAN, Daniel (eds.). *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press, 2002.

GUTHRIE C., RACHLINSKI JJ, WISTRICH AJ. Blinking on the bench: How judges decide cases. In: *Cornell Law Rev*, vol. 93, 2007, p. 1-43.

HAACK, Susan. *Evidence Matters*. New York: Cambridge University Press, 2014.

HAY, Bruce; SPIER, Kathryn. Burdens of Proof in Civil Litigation: An Economic Perspective. In: *The Journal of Legal Studies*, vol. 26, 1997 p. 413-431.

JAMES JR., Fleming. Burdens of proof. In: *Va. L. Rev*, vol. 47, 1961, p. 51 e ss.

JOLLS, Christine; SUNSTEIN, Cass; THALER, Richard. A Behavioral Approach to Law and Economics. In: *Stanford Law Review*, vol. 50, 1998, p. 1471-1550.

KAHNEMAN, Daniel; SLOVIC, Paul.; TVERSKY, Amos. *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Pres, 1982.

_____; TVERSKY, Amos. Judgment under uncertainty: Heuristics and biases. In: KAHNEMAN, Daniel, SLOVIC, Paul., TVERSKY, Amos. *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Pres, 1982.

KASSIN, S. M., ELLSWORTH, P. C., SMITH, V. L. The “general acceptance” of psychological research on eyewitness testimony: A survey of the experts. *American Psychologist*, vol. 44, n. 8, 1989, p. 1089-1098.

LALENG, Per. Sienkiewicz v Greif (UK) Ltd and Willmore v Knowsley Metropolitan Borough Council: A Material Contribution to uncertainty? In: *The Modern Law Review*, vol. 74, 2011, p. 777-793.

LAUDAN, Larry. *Truth, Error and Criminal Law: An Essay in Legal Epistemology*. Cambridge: Cambridge University Press, 2008.

LEACH, Amy-May; CUTLER, Brian L.; VAN WALLENDIAEL, Lori. Lineups and Eyewitness Identification. In: *Annual Review of Law and Social Science*, Vol. 5, 2009, p. 157-178.

LEMPERT, Richard. The New Evidence Scholarship: Analyzing the Process of Proof. In: TILLERS, Peter; GREEN, Eric (orgs.). *Probability and Inference in the Law of Evidence: Uses and Limits of Bayesianism*. London: Kluwer Academic Publishers, 1989, p. 62-102.

LOFTUS, E. F. *Eyewitness testimony*. Cambridge, Massachusetts: Harvard University Press, 1979.

MAGNUSSEN, S.; MELINDER, A.; STRIDBECK, U.; RAJA, A. Q. Beliefs about factors affecting the Reliability of eyewitness testimony: A Comparison of judges, jurors and the general public. In: *Appl. Cognit. Psychol.*, vol. 24, 2010, p. 122-133.

MCCAULIFF, C.M.A. Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees. In: *Vand. L. Rev.*, vol. 35, 1982, p. 1239 e ss.

MEMON, Amina; MASTROBERARDINO, Serena; FRASER, Joanne. Münsterberg’s Legacy: What Does Eyewitness Research Tell Us About the Reliability of Eyewitness Testimony? In: *Appl. Cognit. Psychol.* Vol. 22, 2008, p. 841–851.

MÜNSTERBERG, Hugo. *On the witness stand*. New York: Doubleday, Page & Company, 1908.

PANNINGTON, Nancy; HASTIE, Reid. A cognitive theory of juror decision making: the story model. In: *Cardozo Law Review*, vol. 13, 1991, p. 519 e ss.

PARK, Roger and SAKS, Michael. Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn. In: *Boston College Law Review*, vol. 47, 2005, p. 949–1031.

PICINALI, Federico. The burdens of proof: discriminatory power, weight of evidence, and tenacity of belief. In: *Jurisprudence*, vol. 9, 2018, p. 192-201.

POULTON, E C. *Behavioral Decision Theory: A New Approach*, New York: Cambridge University Press, 1994.

RACHLINSKI, Jeffrey J.; WISTRICH, Andrew J. Judging the Judiciary by the Numbers: Empirical Research on Judges. In: *Annu. Rev. Law Soc. Sci.*, 2017, p. 1-44.

RISINGER, Michael. Innocents Convicted: An Empirical Justified Factual Wrongful Conviction Rate. In: *Journal of Criminal Law and Criminology*, vol. 97, 2007, p. 761 e ss.

SHIN, Hyun Song. The Burden of Proof in a Game of Persuasion. In: *Journal of Economic Theory*, vol. 64, nº 1, 1994, p. 253-64.

SIMON, Rita James, and MAHAN, Linda. Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom. In: *Law & Society Review*, vol. 5, no. 3, 1971, p. 319-330.

STEEL, S. Sienkiewicz v Greif and Exceptional Doctrines of Natural Causation. In: *Journal of European Tort Law*, vol. 2, n. 3, 2011, p. 294-311.

SUSTEIN, Cass. Behavioral law and economics: a progress report. In: *American Law and Economics Review*, vol. 1, 1999, p. 115-157.

TEITELBAUM, Joshua C.; ZEILER, Kathryn (eds.). *Research Handbook on Behavioral Law and Economics*. Cheltenham: Elgar, 2018.

TRIBE, Laurence H. Trial by Mathematics: Precision and Ritual in the Legal Process. In: *Harvard Law Review*, Vol. 84, No. 6, 1971, p. 1329-1393.

WALTON, Douglas N. Burden of Proof. In: *Argumentation* 2, nº 2, 1988, p. 233-54.

WASSERMAN, David. The Morality of Statistical Proof and the Risk of Mistaken Liability. In: *Cardozo L. Rev.*, vol. 13, 1992, p. 935 e ss.

WASSERMAN, Larry. *All of Statistics: A Concise Course in Statistical Inference*. New York: Pittsburgh, Springer, 2003.

WELLS, Gary L. Naked Statistical Evidence of Liability: Is Subjective Probability Enough? In: *Journal of Personality and Social Psychology*, vol. 62, nº 5, 1992, p.739-752.

WHEELAN, Charles. *Naked Statistics: Stripping the Dread from the Data*. New York: W. W. Norton & Company, 2014.

WISTRICH, Andrew; GUTHRIE, Chris; RACHLINSKI, Jeffrey. Can judges ignore inadmissible information? The difficulty of deliberating disregard. In: *University of Philadelphia Law Review*, vol. 153, 2005, p. 1251 e ss.

YAP, Po Jen, Sienkiewicz v Greif: Causation and Risk of Injury. In: *Tort Law Review*, Vol. 19, 2011, p. 61-64.